

基于知识迁移的深度学习无线电信号聚类方法^①

李晓慧^{②*} 陈壮志* 徐东伟* 赵文红** 宣琦*

(* 浙江工业大学信息工程学院网络空间安全研究院 杭州 310023)

(** 嘉兴南湖学院信息工程学院 嘉兴 314001)

摘要 现有的无线电信号调制识别方法在先验数据不足时通常很难对无类标信号进行有效识别。针对这个问题,本文提出了一种基于知识迁移的深度学习无线电信号聚类方法(DTC)。该方法基于样本对比,分析样本间的相似性,并利用卷积神经网络(CNN)提取无线电信号的特征,同时设计了一种预训练框架,通过迁移同领域数据集的知识,有效提升了CNN特征提取能力,实现了引导聚类方向、提升聚类性能的目标。实验结果表明,该方法在多个公开数据集上的聚类性能都显著优于现有的聚类方法。与现有方法相比,DTC在RML 2016.10A和RML 2016.04C数据集上的聚类精度分别提升了30.34%和28.04%。

关键词 信号聚类;深度学习;调制识别;迁移学习;卷积神经网络(CNN)

0 引言

随着现代电磁环境日趋复杂,信号调制识别已成为实现频谱资源监管的重要手段。近年来,深度学习作为一种有效且强大的分类工具,被广泛应用于信号调制识别^[1-6]。然而现有研究大多基于监督学习和半监督学习^[7-9],主要存在以下2个缺点:(1)需要标记样本训练深度学习模型;(2)只能检测训练集所包含的固定类型的信号,一旦先验数据不足或信号的工作状态、信道环境发生变化,很可能导致识别失效^[10]。因此,无线电信号的无监督学习是一个急需解决的难题。

作为一种无监督学习方法,聚类通过捕捉信号间的相关性,将其分为多个簇。由于深度学习模型强大的特征提取能力,深度聚类逐渐兴起,例如基于学生t分布对比的深度嵌入聚类(deep embedded clustering, DEC)^[11]、基于凝聚聚类的联合无监督学习(joint unsupervised learning, JULE)^[12]、基于样本对比的深度自适应图像聚类(deep adaptive image

clustering, DAC)^[13]和基于密度聚类的深度密度图像聚类(deep density-based image clustering, DDC)^[14]等。然而对于无线电信号,直接的聚类很难达到期望的效果。因为从波形上看,信号调制类型相关的特征并不显著,这会导致深度学习模型在聚类初期提取的特征毫无聚集性,模型优化的方向不明确,最终导致无法聚类。目前聚类在无线电信号调制中主要用来分析各种统计特征的重要性^[15]和重构星座图的聚类中心向量^[16-19]。

迁移学习^[20-25]通过对模型参数、样本知识等的迁移利用,在目标数据集样本不足时,可以有效整合利用辅助数据集提供的知识,提升算法性能。因此,通过迁移学习机制弥补深度学习模型在聚类初期特征提取能力不足的缺陷,保证模型能提取到有用的、疏密不同的特征,是提升聚类性能的可行思路。

本文提出了一种用于调制识别的无线电信号深度迁移聚类框架,将特征学习和聚类联合训练,并采用迁移学习策略来增强特征提取能力。该框架的流程如下。首先,使用相同或相关领域的有标记数据

① 国家自然科学基金(61973273)和浙江省自然科学基金(LR19F030001)资助项目。

② 女,1997年生,硕士生;研究方向:信号聚类;联系人,E-mail: 2111903122@zjut.edu.cn。

(收稿日期:2022-04-21)

集对卷积神经网络(convolutional neural network, CNN)进行预训练,由于预训练使用的是有类标的信号,所以可以保证模型所学知识的正确性;然后,使用预训练阶段得到的 CNN 模型对目标数据集中的信号进行初步特征提取并聚类,选择聚类结果中高置信度的信号进行软标注;再根据这些样本进一步微调 CNN 模型;最后重复聚类和微调过程,迭代训练 CNN,直到聚类精度稳定。由于每次模型训练均使用高置信度的信号,模型抗干扰能力强,训练效率高。聚类过程和特征学习过程联合训练,可学到更符合信号聚类需求的隐藏特征。本文的主要贡献可总结如下。

(1)本文提出了一种无监督的无线电信号聚类方法:深度迁移无线电信号聚类(deep transfer clustering of radio signals, DTC)。基于样本对比的思想,分析信号间的相似性,并将特征学习与聚类联合训练,实现无线电信号的聚类,协助识别未知信号的调制类型。

(2)本文设计了一种迁移学习机制,通过对同领域数据集知识的迁移,有效提高 CNN 的特征提取能力,大幅提高信号聚类的性能。

(3)本文将迁移学习与深度聚类结合,实现了

目标属性聚类。通过控制辅助数据集提供的知识,增加特征向量中与调制类型属性相关的特征所占比重,实现对聚类方向的控制,确保样本聚类依据的主要属性为调制类型。

1 深度迁移聚类

本节旨在介绍无线电信号的深度迁移聚类方法。假设 $O = \{x_i\}_{i=1}^n$ 表示无标记的信号数据集,其中 n 为样本个数, x_i 表示第 i 个信号。唯一可知的信息是数据集包含 k 类调制信号,而具体调制类型未知。DTC 将分析这些信号在调制类型方面的相似性,将 O 分成 k 个簇。

DTC 的总体框架如图 1 所示,使用 CNN 作为学习模型来实现特征映射。与其他聚类方法不同的是,为了提高 CNN 对调制信息的特征提取能力,DTC 引入了迁移学习的参数迁移,用已有调制类型标签的数据集作为辅助数据集,将辅助数据集上训练得到的 CNN 参数应用到目标数据集上,辅助数据集与目标数据集互相独立,数据分布不同。DTC 整个框架包括 3 个阶段:数据对齐、预训练和微调聚类。

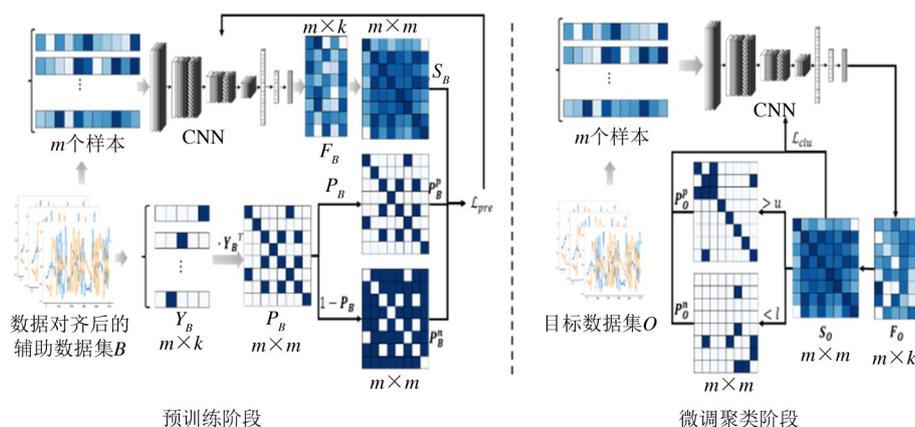


图 1 DTC 的整体框架

1.1 数据对齐

数据对齐的目的是为了实现辅助数据集和待聚类的目标数据集之间信号样本长度的一致性。

辅助数据集 B 是独立于 O 的有标记信号数据集。在聚类之前,使用 B 对模型进行预训练。由于 O 和 B 是相互独立的,2 个数据集的样本长度可能

不同。当 B 中的信号比 O 中的信号短时,通过复制将其扩展至目标长度。例如,长度为 3 的信号“abb”扩展为长度为 8 的“abbabbab”。当 B 中的信号长于 O 中的信号时,基于等间隔采样对其进行压缩,以保留信号整体的波形特征。例如,长度为 8 的信号“abcdefgh”缩短为长度为 3 的“adg”。需要注

意的是,在这里只调整了 B 中信号的长度, O 中的信号会被直接输入到 CNN 中,无需任何调整。

两个数据集中的类别数(调制类型的种类数量)可能存在差异。当 B 中的类别数大于 O 的类别数 k 时,随机选择 k 类的样本用于模型的预训练;当 B 的类别数不够时,所有样本都将被利用,该情况下预训练过程的效果会略有降低。无论辅助数据集类别数是多少,CNN 输出的特征向量的维数始终与 O 的类别数保持一致。

1.2 预训练

数据对齐后,使用 B 中的信号及标签预训练 CNN。为了降低计算量, B 的训练集中的样本被随机划分成多个批次,每个批次包含 m 个信号。对于一个批次的输入信号,CNN 输出 m 个 k 维的特征向量。根据余弦距离,信号 \mathbf{x}_i 和 \mathbf{x}_j 之间的相似度定义为

$$\text{sim}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\| \cdot \|\mathbf{v}_j\|} = \mathbf{f}_i \cdot \mathbf{f}_j \quad (1)$$

其中, \mathbf{v}_i 是信号 \mathbf{x}_i 的特征向量, \mathbf{f}_i 是对 \mathbf{v}_i 归一化后的结果,维度为 k 。所以该批次信号的相似度矩阵:

$$\mathbf{S}_B = \mathbf{F}_B \mathbf{F}_B^T \quad (2)$$

其中, $\mathbf{F}_B = \{\mathbf{f}_i\}_{i=1}^m, \mathbf{S}_B(i, j) = \mathbf{f}_i \cdot \mathbf{f}_j$, 表示信号 \mathbf{x}_i 和 \mathbf{x}_j 之间的相似度, \mathbf{S}_B 的维度为 $m \times m$ 。

然后,将信号的标签转换成维度为 k 的 one-hot 向量。每批信号的标签向量可以组成 $m \times k$ 维的标签矩阵 $\mathbf{Y}_B = \{\mathbf{y}_i\}_{i=1}^m$ 。由此可得二元判断矩阵:

$$\mathbf{P}_B = \mathbf{Y}_B \mathbf{Y}_B^T \quad (3)$$

其中, \mathbf{P}_B 是一个 $m \times m$ 维的布尔矩阵。如果信号 \mathbf{x}_i 和 \mathbf{x}_j 属于同一类别,则矩阵元素 $P_B(i, j) = 1$, 否则 $P_B(i, j) = 0$ 。根据 \mathbf{P}_B , 得到正矩阵 $\mathbf{P}_B^p = \mathbf{P}_B$ 和负矩阵 $\mathbf{P}_B^n = 1 - \mathbf{P}_B$ 。根据二分类交叉熵,预训练过程中的损失函数可以定义为

$$L_{\text{pre}} = -\mathbf{P}_B^p \cdot \log \mathbf{S}_B - \lambda \mathbf{P}_B^n \cdot \log(1 - \mathbf{S}_B) \quad (4)$$

其中,超参数 λ 用于调整正损失和负损失的比例。损失函数用来衡量预测的相似度矩阵与信号真实标签矩阵之间的差异。通过预训练阶段的损失函数,CNN 可以学习信号调制类型相关的知识,提高特征提取能力,并通过余弦相似度判断 2 个信号的调制类型是否相同。这种样本对比的思想类似于 Sim-

CLR^[26],但最大的不同在于 SimCLR 比较的是同一样本的 2 个不同增强版本并学习其特征表示。而本文方法比较的是不同的样本,侧重于总结样本的相似性。

DTC 使用辅助数据集 B 的训练集对 CNN 进行迭代预训练,在每次迭代后,使用验证集评估 CNN 的性能。当验证集上的损失值稳定时,停止预训练,保存 CNN 的参数。预训练使 CNN 具备了一定的特征提取能力,且辅助数据集提供的知识促使 CNN 在提取特征时,更关注与调制类型有关的特征。

1.3 微调与聚类

将预训练得到的 CNN 迁移到目标数据集上,根据目标数据集 O 对 CNN 进行微调,并输出 O 的聚类结果。目标数据集 O 是无标签的。如图 1 右侧所示,同样,将目标数据集 O 分批输入 CNN,批次大小与预训练阶段相同。对于每个批次的输入信号,CNN 输出信号的特征矩阵为 $\mathbf{F}_O, \mathbf{F}_O$ 的维度为 $m \times k$ 。根据 \mathbf{F}_O 得到该批信号的相似度矩阵:

$$\mathbf{S}_O = \mathbf{F}_O \mathbf{F}_O^T \quad (5)$$

其中, \mathbf{S}_O 的维度为 $m \times m$ 。由于 O 中的样本是无标注的,DTC 使用软标签代替样本的真实标签。将 \mathbf{S}_O 的元素分别与上限阈值 u 和下限阈值 l 进行比较,以衡量 2 个信号属于同一簇的可能性。相似度大于上限阈值和小于下限阈值的样本对将被筛选用于 CNN 的微调训练,不在此范围内的样本对暂时忽略。

根据相似度矩阵构造正矩阵 \mathbf{P}_O^p 和负矩阵 \mathbf{P}_O^n , 矩阵元素定义为

$$P_O^p(i, j) = \begin{cases} 1 & S_O(i, j) \geq u \\ 0 & S_O(i, j) < u \end{cases} \quad i, j = 1, \dots, m \quad (6)$$

$$P_O^n(i, j) = \begin{cases} 0 & S_O(i, j) \geq l \\ 1 & S_O(i, j) < l \end{cases} \quad i, j = 1, \dots, m \quad (7)$$

如果 $P_O^p(i, j) = 1$ 表示信号 \mathbf{x}_i 和 \mathbf{x}_j 被暂时标记为同类信号, $P_O^n(i, j) = 1$ 则表示 \mathbf{x}_i 和 \mathbf{x}_j 被暂时标记为不同类信号。 \mathbf{P}_O^p 的目的是选择具有高置信度的同类别样本对,而 \mathbf{P}_O^n 的目的是选择高置信度的不同类别样本对,聚类训练阶段的损失函数定义为

$$L_{\text{cla}} = -P_o^y \cdot \log S_o - \lambda P_o^x \cdot \log(1 - S_o) \quad (8)$$

每次训练中,只有当前置信度较高的样本对被用于训练 CNN,因为高置信度样本的软标签是正确的可能性更大,也意味着这些样本在调制类型上的特征更鲜明,有利于 CNN 学习表示 O 中样本的调制类型特征。随着微调的进行,由于高置信度样本的标签正确性相对较高,模型对当前数据集的特征提取能力得到了加强,所以会出现更多的高置信度样本参与微调,最终提高模型性能。在训练过程中,由于 DTC 采用余弦距离衡量特征向量间的相似程度,所以来自不同类别的信号的特征向量会趋于相互垂直。特征向量的维度为 k ,与类别数相同。随着训练的进行,经过归一化处理的特征向量会趋于 one-hot 向量的形式。归一化后的特征向量实际代表信号在各个类别中的概率分布。

CNN 的结构如图 2 所示,包括 4 层二维卷积层 (Conv) 和 2 层全连接层 (Dense),每层都使用整流线性 (rectified linear unit, ReLU)^[27] 激活函数,并在 ReLU 层之前增加了批标准化 (batch normalization, BN)^[28] 以防止过拟合。同时,BN 层可以将数据的分布调整为正态分布,在每次输入数据的分布不同时保证模型的泛化性能。为了去除冗余信息,在第 2 和第 3 卷积层的 ReLU 函数之后添加了最大池化层 (MaxPool)。最大池化层的输出后也添加了 BN 层。CNN 的 4 个卷积层分别包含 32、128、128 和 32 个过滤器。最后 2 个全连接层分别包含 64 和 k 个神经元。模型的最后使用 softmax 分类器输出概率分布。CNN 的更多详细参数设置参见表 1。本文所有实验均基于 TensorFlow 深度学习框架实现,并使用 Adam 优化器对模型进行训练。

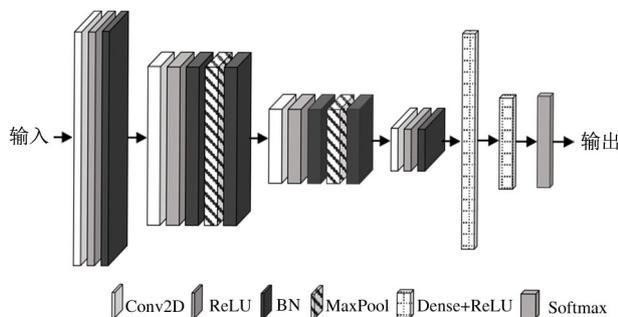


图 2 DTC 使用的 CNN 结构

表 1 CNN 的结构设置

层	核大小	步长	滤波器数量
Conv 1	7×1	[1, 1]	32
Conv 2	9×2	[1, 1]	128
MaxPool 1	8×1	[2, 1]	—
Conv 3	9×1	[1, 1]	128
MaxPool 2	8×1	[2, 1]	—
Conv 4	9×1	[1, 1]	32
Dense 1	—	—	64
Dense 2	—	—	k

2 整体实验设置

2.1 数据集

本文使用 3 个公开数据集: RML 2016. 10A、RML 2016. 04C 和 RML 2018. 01A。

RML 2016. 10A 是一个使用 GNU Radio 生成的公共数据集。该数据集包含 11 种调制类别的样本,包括数字调制的 BPSK、QPSK、8PSK、16QAM、64QAM、BFSK、CPFSK 和 PAM4,以及模拟调制的 WB-FM、AM-SSB 和 AM-DSB。数据集包含从 $-20 \sim 18$ dB 的 20 种不同信噪比的信号样本,每个样本的形状为 128×2 。

RML 2016. 04C 是一个使用 GNU Radio 生成的公共合成数据集,由 11 种不同调制类型的信号组成。样本的调制类别种类、信噪比和形状与 RML 2016. 10A 相同,但该数据集的生成过程更接近真实信号。

RML2018. 01A 是一个包括 24 种数字和模拟调制类型的数据集,具体调制类型有: 32PSK、16APSK、32QAM、FM、GMSK、32APSK、OQPSK、8ASK、BPSK、8PSK、AM-SSB-SC、4ASK、16PSK、64APSK、128QAM、128APSK、AM-DSB-SC、AM-SSB-WC、64QAM、QPSK、256QAM、AM-DSB-WC、OOK、16QAM。每个样本的形状为 1024×2 。

由于高信噪比的信号受噪声影响更小,为了突出显示各种方法的效果,在本文的实验中,只使用了信噪比最高的信号。实验中使用的数据集的详细信息如表 2 所示。

表 2 使用的信号数据集

数据集	训练集大小	测试集大小	信噪比	类别数
RML 2016. 10A	8800	2200	18 dB	11
RML 2016. 04C	6459	1644	18 dB	11
RML 2018. 01A	78 643	19 661	30 dB	24

2.2 对比方法

为了说明 DTC 的性能,本文选择了传统聚类方法 K-means 和几种深度聚类方法与 DTC 进行对比。由于信号领域尚未有成熟的无监督学习的方法,本文选择了近年来提出的 2 种通用性较强的高性能深度聚类方法 DEC 和 DAC。同时作为参考,本文列出了 2 种半监督方法的结果,包括基于生成对抗网络的信号分类方法(signal classifier generative adversarial networks, SCGAN)^[9]和半监督信号识别卷积神经网络(semi-supervised signal recognition convolutional neural network, SSRCNN)^[8]。

2.3 评价指标

本文选择 3 种经典聚类指标来评估实验结果,包括聚类精度(accuracy, ACC)、调整兰德系数(adjusted rand index, ARI)和归一化互信息(normalized mutual information, NMI)。指标数值越高表明聚类性能越好。

2.4 超参数

由于该模型对超参数 λ 并不敏感,当 λ 发生量级程度的改变时,才会对实验结果产生影响。因此可以对 λ 取 0.01、0.1、1 等不同数量级的值进行实验,根据实验结果确定 λ 的取值。在本文实验中,预训练阶段设置 $\lambda = 0.1$,微调聚类阶段设置 $\lambda = 100$ 。对于上下限阈值,本文实验以 0.1 为间隔,在 $[0, 1]$ 范围内为 l 取不同的值,在 $[l, 1]$ 范围内为 u 取不同的值,经过 66 次测试,最终确定上限阈值 $u = 0.95$,下限阈值 $l = 0.455$ 。

3 实验结果

3.1 DTC 的聚类实验

本文在 3 个数据集上进行交叉聚类实验。每次实验选择 2 个数据集,分别作为目标数据集和辅助数据集。由于 RML 2018. 01A 包含 24 种类型的样

本,其他 2 个数据集只有 11 类。当使用 RML 2018. 01A 作为辅助数据集时,仅使用其中 11 类样本进行预训练。目标数据集和辅助数据集在调制类型上的重叠可能会带来过好的提升效果,干扰 DTC 的性能评估。为了消除这个因素的干扰,RML 2018. 01A 作为辅助数据集时,使用的是 11 种与 RML 2016. 10A 调制类型完全不重合的信号样本。所选样本类型为 32QAM、FM、GMSK、32APSK、OQPSK、8ASK、4ASK、16PSK、64APSK。

DTC 在预训练阶段仅使用辅助数据集的训练集训练 CNN,在微调聚类阶段使用目标数据集的所有样本(包括训练集和测试集)微调 CNN 模型。训练过程和微调过程的批次大小、学习率和训练周期数相同。批量大小 m 和迭代训练次数都是 150。对于不同的训练周期,学习率的值不同。前 20 个周期的学习率为 0.001,第 21 ~ 50 个周期的学习率为 0.0008,最后 100 个周期的学习率为 0.0005。训练过程中,损失值会持续减少直到稳定,此时停止训练并保存模型参数。

实验结果如表 3 所示,第 1 列显示了所有实验的预训练阶段使用的辅助数据集,第 1 行是相应的目标数据集。表中也列出了 DTC 在没有预训练的情况下的聚类结果以用于对比。实验结果表明,利用辅助数据集对 CNN 进行预训练后,DTC 的性能具有显著提升。在任何一组实验中,聚类结果的 3 个评估指标都有明显上升。以 RML2016. 10A 数据集上的实验结果为例,未经预训练时,DTC 的聚类指标 NMI、ARI、ACC 分别为 0.4045、0.2693、0.3809。然而,在使用数据集 RML2016. 04C 进行预训练后,3 个指标分别增加到 0.6780、0.5445 和 0.6552。表 3 显示,在 RML 2016. 10A 数据集的聚类结果中,使用 RML 2016. 04C 作为辅助数据集时,DTC 的性能提升最显著。而 RML 2016. 04C 数据集的聚类结果中,使用 RML 2016. 10A 作为辅助数据集时,DTC 的性能提升最显著。原因在于数据集 RML 2016. 10A 和 RML 2016. 04C 的相似度较高,两者的调制类型和信号长度完全相同。值得注意的是,即使数据集 RML 2018. 01A 在调制类型和信号长度上与另外 2 个数据集完全不同,选择 RML 2018. 01A 作为辅助

数据集时,仍然可以提高另外 2 个数据集的聚类性能。同样,RML 2018.01A 作为目标数据集时,其他 2 个数据集预训练得到的 CNN 参数也能提升 RML

2018.01A 上的聚类性能。这说明 DTC 具有一定的泛化能力。

表 3 3 个数据集上交叉预训练的聚类结果

目标数据集 评价指标	RML2016.10A			RML2016.04C			RML2018.01A		
	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC
无预训练	0.4045	0.2693	0.3809	0.4880	0.3020	0.4830	0.3908	0.2229	0.2662
辅助数据集	RML 2016.10A	——	——	0.7852	0.6431	0.7023	0.5533	0.2873	0.3027
	RML 2016.04C	0.6780	0.5445	0.6552	——	——	0.5808	0.3043	0.3276
	RML 2018.01A	0.6213	0.4426	0.5640	0.6984	0.5040	0.5960	——	——

3.2 对预训练阶段的消融研究

3.1 节中的实验证明了预训练可以提高模型的聚类性能。为了研究预训练对最终结果的影响,本文进行了消融实验。取原辅助数据集训练集的 5%、10%、20%、30%、40% 和 50% 的样本进行预训练,分别进行聚类实验。样本挑选采用分层随机抽取策略,每种调制类型的样本所占比例与原训练集中的一致。表 4 展示了 6 组交叉实验(即 exp 1~6)中辅助数据集和目标数据集的选取情况。结果如图 3 所示。可以看出,随着辅助数据数量的减少,目标数据集上的聚类精度逐渐下降,但在辅助数据集的样本数量下降至 10% 前,精度总体下降幅度不大。部分实验结果中,样本数量降低,精度出现大幅波动,原因可能与辅助样本提供的知识量有关。辅助数据数量减少,提供的调制信息减少,所以聚类精度出现下降。但样本提供的调制信息存在重叠,所以当辅助数据数量足够充足时,即使数据量大幅减少,精度下降幅度仍然不大。在部分情况下,若辅助数据集中不同样本提供的信息间重叠程度相对较

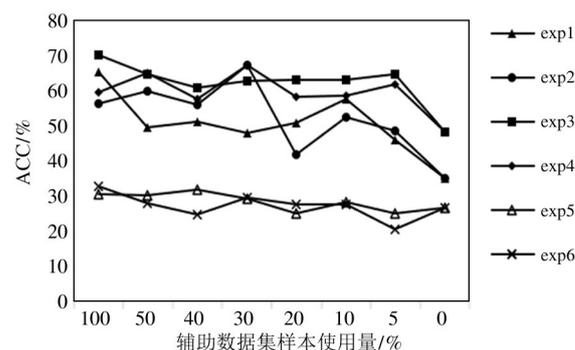


图 3 消融实验结果

小,信息总量较高,就会产生精度较高的实验结果;若部分样本提供的信息过少,也可能对聚类产生不利影响。

3.3 信号调制种类的影响

为了验证 DTC 在不同调制类型上的普适性,本文进行了以下实验。为了消除类别数不一致产生的干扰,保证辅助数据集的类别数量与目标数据集相同,本文在每个数据集中随机选择了 10 类样本进行实验,并将每个数据集的 10 类样本等分为第 1 部分和第 2 部分。划分原则为任何数据集的第 1 部分不与其他数据集的第 2 部分调制类型重叠。划分后的各部分数据集的具体调制信息见表 5。

此项实验使用各数据集的第 1 部分进行预训练,在其他数据集的第 2 部分上进行微调聚类,结果如表 6 所示。实验结果表明,即使辅助数据集和目标数据集的调制类型完全不同,DTC 也能有很好的聚类效果。无预训练时,DTC 在 RML 2016.10A 数据集上的聚类精度为 51.62%。经过预训练,聚类精度提高至 99.52%,其他性能指标也有显著提升。

表 4 exp 1~6 中的辅助数据集和目标数据集

	辅助数据集	目标数据集
exp 1	RML2016.04C	RML2016.10A
exp 2	RML2018.01A	RML2016.10A
exp 3	RML2016.10A	RML2016.04C
exp 4	RML2018.01A	RML2016.04C
exp 5	RML2016.10A	RML2018.01A
exp 6	RML2016.04C	RML2018.01A

表 5 对 3 个数据集的划分结果

数据集	第 1 部分	第 2 部分
RML 2016. 10A	WBFM, QPSK, 64QAM, 16QAM, 4PAM	CFSK, CPFSK, BPSK, 8PSK, AM-SSB
RML 2016. 04C	WBFM, QPSK, 64QAM, 16QAM, 4PAM	CFSK, CPFSK, BPSK, 8PSK, AM-SSB
RML 2018. 01A	32PSK, 16APSK, 32QAM, FM, GMSK	32APSK, OQPSK, 8ASK, 16PSK, 64APSK

表 6 3 个数据集的交叉实验的聚类结果(5 类)

目标数据集 评价指标	RML2016. 10A			RML2016. 04C			RML2018. 01A		
	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC
无预训练	0.4266	0.3505	0.5162	0.2863	0.1772	0.4160	0.0005	0.0002	0.2077
辅助数据集									
RML2016. 10A	—	—	—	0.8044	0.7549	0.8667	0.4185	0.2502	0.4122
RML2016. 04C	0.8749	0.8052	0.9952	—	—	—	0.4089	0.2467	0.4049
RML2018. 01A	0.9805	0.9861	0.9944	0.7371	0.6889	0.7819	—	—	—

DTC 在 RML 2016. 04C 数据集上的聚类精度也从 41.60% 提高到了 86.67%。

为了进一步验证实验结果,排除偶然性,本文对 RML 2016. 10A 和 RML 2016. 04C 的 4 个部分进行了交叉实验,实验结果如表 7 所示。为方便表述,RML2016. 10A 的第 1 部分记为 P1,第 2 部分记为 P2;RML 2016. 04C 的第 1 部分记为 P3,第 2 部分记为 P4。P1 和 P3 中样本的调制类型相同,P2 和 P4 的调制类型相同。

从表 7 可以看出,无论辅助数据集和目标数据

集的调制类型是否相同,聚类精度都有所提高。当 2 个数据集的调制类型相同时,预训练对聚类精度的提升最为显著。以 P3 为辅助数据集时,对 P1 的聚类精度可以达到 72%,比次优结果高 21.7%;以 P2 为辅助数据集时,对 P4 的聚类精度可以达到 88%,同样高于其他实验结果。同时,辅助集和目标集的生成方式更相似时,也可能获得更好的性能结果,例如,当设置 P3 为目标数据集时,选择 P4 作为辅助数据集时聚类精度最高。

表 7 RML 2016. 10A 和 RML 2016. 04C 交叉实验的聚类结果(5 类)

目标数据集 评价指标	P1			P2			P3			P4		
	NMI	ARI	ACC									
无预训练	0.5565	0.4847	0.5080	0.7036	0.6040	0.7015	0.3531	0.3097	0.5130	0.3503	0.2347	0.4559
辅助数据集												
P1	—	—	—	0.9844	0.9883	0.9954	0.7940	0.7922	0.7885	0.7562	0.7004	0.7802
P2	0.5359	0.4563	0.5916	—	—	—	0.8025	0.7167	0.7440	0.8737	0.8465	0.8837
P3	0.6372	0.5839	0.7272	0.8749	0.8052	0.9255	—	—	—	0.7762	0.6988	0.7967
P4	0.5163	0.4418	0.5896	0.9855	0.9895	0.9958	0.9386	0.9427	0.9779	—	—	—

3.4 与现有方法的对比

本文将 DTC 与 3.2 节中提到的其他方法进行了比较。对于 DEC 的 DAC,实验中的各种参数设置与原论文的推荐值一致。由于 K-means 和 DEC 所需的输入数据形式为一维,信号的 Q 通道数据被拼接在 I 通道后方,形成一维数据。而在 DAC 的实验中,信号的形状被调整为 $n \times n$ 。SCGAN 和 SSRCNN 的结果来自原论文,都是在有类标数据占比 10% 的情况下,高信噪比信号上的半监督实验结果。

对比结果见表 8。其中“DTC (无预训练)”行是不进行预训练的结果。“DTC”行代表经过预训练的聚类结果,引用的是表 3 中的最优值。实验结果显示 DTC 的聚类效果明显优于所有其他无监督聚类方法。特别是在 RML2016. 10A 上,DTC 的聚类精度达到了 0.6552,比次优方法 DAC 高出 86.2%,比半监督方法 SSRCNN 的精度低 0.04。在 RML2016. 04C 上,DTC 的聚类精度为 0.7023,比 DAC 高出 44.5%,超过了 SCGAN。在 RML2018. 01A 上,DTC 的聚类

精度为 0.3276, 比 DAC 高出 23.3%。这些结果表明, 对于无线电信号无监督学习这一项具有挑战性

的任务, DTC 是一种可行的方法。

表 8 各种方法在 3 个数据集上的聚类结果

目标数据集 评价指标	RML 2016.10A			RML 2016.04C			RML 2018.01A			
	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	
无监督	K-means ^[29]	0.1793	0.0787	0.1897	0.3184	0.1583	0.3119	0.3005	0.0579	0.1112
	DEC ^[11]	0.1780	0.0857	0.1911	0.3944	0.2196	0.4140	0.4393	0.0596	0.1363
	DAC ^[13]	0.3427	0.2275	0.3518	0.4827	0.2890	0.4219	0.4242	0.1520	0.2656
	DTC(无预训练)	0.4045	0.2693	0.3809	0.4880	0.3020	0.4830	0.3908	0.2229	0.2662
	DTC	0.6780	0.5445	0.6552	0.7852	0.6431	0.7023	0.5808	0.3043	0.3276
半监督	SCGAN ^[9]	—	—	—	—	—	0.4600	—	—	—
	SSRCNN ^[8]	—	—	0.6931	—	—	0.9380	—	—	—

4 结 论

借助卷积神经网络强大的特征提取能力, 本文提出了一种新型端到端的无线电信号深度迁移聚类方法(DTC), 将深度学习和迁移学习融合到一个整体框架中, 以此提高聚类性能。实验结果表明, 该方法具有一定的泛化能力, 即使辅助数据集与目标数据集的调制类型不重合, DTC 也能实现较好的聚类效果。与现有方法相比, 本文所提创新方法在 3 个公共无线电信号数据集上取得了更优的性能。

无线电信号的无监督学习的难点在于特征提取。本文的实验证明了利用迁移学习机制提升深度学习模型特征提取能力的可行性。未来的研究方向一方面将尝试改进 DTC 的框架, 以期实现跨领域数据集上的知识迁移, 另一方面将尝试其他增强方法, 进一步提升 DTC 的聚类性能。

参 考 文 献

- [1] CHEN Z, CUI H, XIANG J, et al. SigNet: a novel deep learning framework for radio signal classification [J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(2) : 529-541.
- [2] JIAO J, SUN X, ZHANG Y, et al. Modulation recognition of radio signals based on edge computing and convolutional neural network [J]. Journal of Communications and Information Networks, 2021, 6(3) : 280-300.
- [3] NJOKU J N, MOROCHO-CAYAMCELA M E, LIM W. CGDNet: efficient hybrid deep learning model for robust automatic modulation recognition [J]. IEEE Networking Letters, 2021, 3(2) : 47-51.
- [4] WANG T, HOU Y, ZHANG H, et al. Deep learning based modulation recognition with multi-cue fusion [J]. IEEE Wireless Communications Letters, 2021, 10(8) : 1757-1760.
- [5] LIU K, GAO W, HUANG Q. Automatic modulation recognition based on a DCN-BiLSTM network [J]. Sensors, 2021, 21(5) : 1577.
- [6] 何继爱, 张文启. 通信信号调制识别技术及其发展 [J]. 高技术通讯, 2016, 26(2) : 157-165.
- [7] WANG Y, GUI G, GACANIN H, et al. Transfer learning for semi-supervised automatic modulation classification in ZF-MIMO systems [J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2020, 10(2) : 231-239.
- [8] DONG Y, JIANG X, CHENG L, et al. SSRCNN: a semi-supervised learning framework for signal recognition [J]. IEEE Transactions on Cognitive Communications and Networking, 2021, 7(3) : 780-789.
- [9] LI M, LI O, LIU G, et al. Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks [J]. Sensors, 2018, 18(11) : 3913.
- [10] ZHANG B, JI G. Multi-angle development analysis for automatic modulation classification technology [C] // Proceedings of the 5th International Conference on Electronic Information Technology and Computer Engineering. Xiamen: ICEECE, 2021 : 276-281.
- [11] XIE J, GIRSHICK R, FARHADI A. Unsupervised deep embedding for clustering analysis [C] // International Conference on Machine Learning. New York: ICML, 2016 : 478-487.
- [12] YANG J, PARIKH D, BATRA D. Joint unsupervised learning of deep representations and image clusters [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016 : 5147-5156.
- [13] CHANG J, WANG L, MENG G, et al. Deep adaptive image clustering [C] // Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE,

- 2017:5879-5887.
- [14] REN Y, WANG N, LI M, et al. Deep density-based image clustering [J]. Knowledge-Based Systems, 2020, 197:105841.
- [15] DALDAL N, POLAT K, GUO Y. Classification of multi-carrier digital modulation signals using NCM clustering based feature-weighting method[J]. Computers in Industry, 2019,109:45-58.
- [16] JAJOO G, KUMAR Y, YADAV S K, et al. Blind signal modulation recognition through clustering analysis of constellation signature [J]. Expert Systems with Applications, 2017,90:13-22.
- [17] YANG F, YANG L, WANG D, et al. Method of modulation recognition based on combination algorithm of K-means clustering and grading training SVM[J]. China Communications, 2018,15(12):55-63.
- [18] TIAN J, PEI Y, HUANG Y, et al. Modulation-constrained clustering approach to blind modulation classification for MIMO systems [J]. IEEE Transactions on Cognitive Communications and Networking, 2018,4(4):894-907.
- [19] ZHAO Z, YANG A, GUO P, et al. A density clustering algorithm for simultaneous modulation format identification and OSNR estimation [J]. Applied Sciences, 2020, 10(3):1095.
- [20] DAI W, YANG Q, XUE G, et al. Boosting for transfer learning [C]//Proceedings of the 24th International Conference on Machine Learning. Corvalis: ICML, 2007: 193-200.
- [21] RAINA R, BATTLE A, LEE H, et al. Self-taught learning: transfer learning from unlabeled data [C]//Proceedings of the 24th International Conference on Machine Learning. Corvalis: ICML, 2007:759-766.
- [22] EVGENIOU T, PONTIL M. Regularized multi-task learning [C]//Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle: Association for Computing Machinery, 2004:109-117.
- [23] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks? [EB/OL]. (2014-11-06) [2022-08-17]. <http://arxiv.org/pdf/1411.1792.pdf>.
- [24] 黎传琛, 白勇, 陈益民. 基于迁移学习的火焰图像识别技术研究[J]. 高技术通讯, 2019,29(3):274-282.
- [25] ZHANG X Y, SHI H, LI C, et al. Weakly-supervised action localization via embedding-modeling iterative optimization [J]. Pattern Recognition, 2021,113:107831.
- [26] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations [C]//International Conference on Machine Learning. Vienna: ICML, 2020:1597-1607.
- [27] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [C]//International Conference on Machine Learning. Haifa: ICML, 2010:807-814.
- [28] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//International Conference on Machine Learning. Lille: IEEE, 2015:448-456.
- [29] MACQUEEN J. Some methods for classification and analysis of multivariate observations [C]//Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press, 1967,1(14):281-297.

A deep learning radio signal clustering method based on knowledge transfer

LI Xiaohui*, CHEN Zhuangzhi*, XU Dongwei*, ZHAO Wenhong**, XUAN Qi*

(* Institute of Cyberspace Security, College of Information Engineering,
Zhejiang University of Technology, Hangzhou 310023)

(** School of Information Engineering, Jiaying Nanhu University, Jiaying 314001)

Abstract

The existing radio signal modulation identification methods are usually difficult to effectively identify the unclassified signal when the prior data is insufficient. To solve this problem, this paper proposes a deep transfer clustering (DTC) of radio signals method based on knowledge transfer. This method analyzes the similarity between samples based on sample comparison, and uses a convolutional neural network (CNN) to extract the features of radio signals. At the same time, a pre-training framework is designed, which effectively improves the feature extraction ability of CNN by transferring the knowledge of the same domain dataset and achieves the goal of guiding the clustering direction and improving the clustering performance. The experimental results show that the clustering performance of this method is significantly better than the existing clustering methods on multiple public datasets. Compared with existing methods, the clustering accuracy of DTC on the RML2016.10A and RML2016.04C datasets is improved by 30.34% and 28.04%, respectively.

Key words: signal clustering, deep learning, modulation recognition, transfer learning, convolutional neural network (CNN)